

COMMUNICABILITY ANGLE AND THE SPATIAL EFFICIENCY OF NETWORKS

ERNESTO ESTRADA[†] AND NAOMICHI HATANO[‡]

Abstract. We introduce the concept of communicability angle between a pair of nodes in a graph. We provide strong analytical and empirical evidence that the average communicability angle for a given network accounts for its spatial efficiency on the basis of the communications among the nodes in a network. We determine characteristics of the spatial efficiency of more than a hundred real-world complex networks that represent complex systems arising in a diverse set of scenarios. In particular, we find that the communicability angle correlates very well with the experimentally measured the relative packing efficiency of proteins that are represented as residue networks. We finally show how we can modulate the spatial efficiency of a network by tuning the weights of the edges of the networks. This allows us to predict effects of external stresses on the spatial efficiency of a network as well as to design strategies to improve important parameters in real-world complex systems.

Key words. complex network; communicability; graph distance; graph planarity; Euclidean distance

AMS subject classifications. 05C12; 05C50; 05C82; 05C10

1. Introduction. Graphs are frequently used to represent discrete objects both in abstract mathematics and computer sciences as well as in applications, such as theoretical physics, biology, ecology and social sciences [24, 13]. In the particular case of representing the networked skeleton of complex systems, graphs receive the denomination of complex networks; we will hereafter use graphs and networks interchangeably.

The complex networks are ubiquitous in many real-world scenarios, ranging from the biomolecular — those representing gene transcription, protein interactions, and metabolic reactions — to the social and infrastructural organization of modern society [10, 30, 8]. In many of these networks, nodes and edges are used to represent physically embedded objects [3], namely *spatial networks*. In urban street networks, for instance, the nodes describe the intersection of streets, which are represented by the edges of the graph. These streets and their intersections are embedded in the two-dimensional space representing the surface occupied by the corresponding city [26]. Another spatial network is the brain networks, in which the nodes account for brain regions embedded in the three-dimensional space occupied by the brain, while the edges represent the communication or physical connections among these regions [6]. We can also capture the three-dimensional structure of proteins by means of the residue networks in which nodes describe amino acids and the edges represent physical interactions among them. Other examples include the following: infrastructures, such as the Internet, transportation networks, water and electricity supply networks, *etc.* [3]; anatomical networks, such as vascular and organ/tissue networks; the networks of channels in fractured rocks; the networks representing the corridors and galleries in animal nests; even for others see Ref. [10] and references therein.

A natural question that arises in the analysis of spatial networks is how efficiently they use the available geographical space in which they are embedded. In a protein, for instance, the linear polypeptide chain is very much folded into a three-

[†]Department of Mathematics & Statistics, University of Strathclyde, Glasgow G11XQ, UK

[‡]Institute of Industrial Science, University of Tokyo, Komaba, Meguro, Tokyo 153-8505, Japan

dimensional space in order to minimize the volume occupied inside the cell [9]. In airport transportation networks the nodes are embedded into the two-dimensional space represented by the surface of a country or continent, but the connections between the airports occupy the available three-dimensional space (it might be argued that they use a four-dimensional space as two flights can intersect in space but at different times), which increases the spatial efficiency of these networks. In contrast, the planarity of urban street networks [7], that is, the fact that we can draw the network in a plane without any intersection of the edges, implies that both nodes and edges are embedded in a two-dimensional space, which in general decreases the number of alternative routes between different points in the network. This relatively poor spatial efficiency of modern cities, *i.e.*, the non-existence of three-dimensional cities (although they have been already planned; see Chapter 3 in Ref. [10] and references therein), has posed a serious challenge to their continuous growth in view of their threat to the natural environment. Although the planarity may be an important part of this problem, it is definitively not the only one. Two planar networks, *e.g.*, two cities, can display significantly different spatial efficiency, and the same is true for pairs of non-planar networks.

The concept of *spatial efficiency* is adapted here from economics, where it is frequently used to describe how much time, effort and cost a given arrangement produces for governments, businesses and households to conduct their activities as compared to alternative arrangements; see Ref. [32] and references therein. This concept has a lot to do with the efficiency in communication among the parts of the system under study and as so it is a well-posed problem for its analysis beyond spatial networks.

In this context of communication among the nodes of a network, we [15] have introduced the *communicability function* as a way to quantify how much information can flow from one node to another in a network; see also Refs. [16, 17]. We regard the quantity G_{pq} , which we will define in Eq. (2.2) below, as the amount of information that departs from a node p and arrives at a node q . On the other hand, we regard G_{pp} as the amount of information that departs from the original node p and never arrives at the destination q , because it is returned to its originator. Let us call the first amount of information the successful information and the second the frustrated one. Then, the goodness of communication between the two nodes is given by a combination of the successful to the frustrated amount of information. Increasing the amount of successful information and reducing the amount of frustrated one improves the quality of communication between the two nodes. This has lead to the definition of a quantity [11, 12, 19] that has been proved to be a distance between two nodes.

In the present paper, we show a remarkable mapping of each node of a network to a point on the surface of a hypersphere. We prove that the distance defined based on the communicability function is indeed the cord distance between the two points on the hypersphere. We can thereby assign a Euclidean angle to each pair of nodes which represents the communication efficiency between them. We then analyze various networks using the angle, which we refer to as the communicability angle hereafter, and provide evidence that this angle accounts for the spatial efficiency of networks.

2. Preliminaries. In this section we shall present some of the definitions, notations, and properties associated with networks to make this work self-contained. A *graph* $\Gamma = (V, E)$ is defined by a set of n nodes (vertices) V and a set of m edges (links) $E = \{(p, q) | p, q \in V\}$ between the nodes. An edge is said to be *incident* to a vertex p if there exists a node $q (\neq p)$ such that either $(p, q) \in E$ or $(q, p) \in E$. The *degree* of a vertex, denoted by k_p , is the number of edges incident to p in Γ . The

graph is said to be *undirected* if the edges are formed by unordered pairs of vertices. A *walk* of length ℓ in Γ is a set of nodes $p_1, p_2, \dots, p_\ell, p_{\ell+1}$ such that for all $1 \leq i \leq \ell$, $(p_i, p_{i+1}) \in E$. A *closed walk* is a walk for which $p_1 = p_{\ell+1}$. A *path* is a walk with no repeated nodes. A graph is *connected* if there is a path connecting every pair of nodes. A graph with unweighted edges, no self-loops (edges from a node to itself), and no multiple edges is said to be *simple*. Throughout this work, we will always consider undirected, simple, and connected networks.

More specifically, we will consider graphs which are defined as follow. The path graph P_n is a graph with n nodes, $n-2$ of which have degree 2 and the remaining two have degree 1. The complete graph K_n is the graph with n nodes and $n(n-1)/2$ edges. The complete bipartite graph K_{n_1, n_2} is the graph with $n = n_1 + n_2$ nodes split into two disjoint sets, one containing n_1 nodes and the other containing n_2 nodes, while the edges connect every node in one set with every one in the other. The particular case $K_{1, n-1}$ is known as the star graph. A graph is planar if it can be drawn in a plane without any edges crossing. The following is a well-known characterization of the planar graphs known as the Kuratowski theorem (see Ref. [23]).

THEOREM 2.1. *A network is planar if and only if it has no subgraph homeomorphic to K_5 or $K_{3,3}$.*

Let us consider a matrix A called the adjacency matrix, whose elements are $A_{pq} = 1$ if $(p, q) \in E$ and zero otherwise. For undirected simple finite graphs, A is a real symmetric matrix. We can therefore decompose it into the form

$$(2.1) \quad A = U \Lambda U^T,$$

where Λ is a diagonal matrix containing the eigenvalues of A , which we label in non-increasing order $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, and $U = [\vec{\psi}_1, \dots, \vec{\psi}_n]$ is an orthogonal matrix, where $\vec{\psi}_\mu$ is an eigenvector associated with λ_μ . Because we consider connected graphs, A is irreducible; the Perron-Frobenius theorem then dictates that $\lambda_1 > \lambda_2$ and that we can choose $\vec{\psi}_1$ such that its components $\psi_1(p)$ are positive for all $p \in V$.

An important quantity for studying communication processes in networks is the communicability function [15, 17, 16], defined for each pair of nodes p and q as

$$(2.2) \quad G_{pq} = \sum_{k=0}^{\infty} \frac{(A^k)_{pq}}{k!} = (e^A)_{pq} = \sum_{k=1}^n e^{\lambda_k} \vec{\psi}_k(u) \vec{\psi}_k(v).$$

The factor $(A^k)_{pq}$ counts the number of walks of length k starting at the node p and ending at the node q . The communicability function is a sum of the numbers, each weighted by the factor $1/k!$ so that shorter walks may be more influential than longer ones. The importance of this function lies in the fact that it takes account of long walks too; even two nodes connected by a very long shortest path can have a strong communication if they are connected by very many longer walks. The diagonal term G_{pp} characterizes the degree of participation of the node p in all subgraphs of the network. It is thus known as the subgraph centrality of the corresponding node [18].

It is possible to define several distance measures on networks. The most common one is the *shortest-path* or *geodesic distance* between two nodes $p, q \in V$, which is defined as the length of the shortest path connecting these nodes. We will write $d(p, q)$ to denote the distance between p and q . Here we will refer to the average of the shortest path distance in the graph as the average path length, as usual in network theory. Another distance among the nodes of a graph is the so-called resistance

distance [27] which is defined by $\Omega_{pq} = L_{pp}^+ + L_{qq}^+ - 2L_{pq}^+$, where L^+ is the Moore-Penrose pseudoinverse of the Laplacian matrix of the network [34, 21]; the network Laplacian is defined by $L = K - A$ with $K = \text{diag}(k_i)$.

In the next section we will introduce a third distance defined recently on the basis of the communicability function. It is novel in the sense that longer walks than the shortest path are taken into account.

3. Communicability distance. The new distance function is defined as [11, 12]

$$(3.1) \quad \xi_{pq}^2 = G_{pp} + G_{qq} - 2G_{pq},$$

which we will refer to as the communicability distance between the nodes p and q in Γ . The intuition behind it is that when two nodes p and q communicate with each other, the quality of their communication depends on two factors: (i) how much information departing from the node p (q) arrives at the node q (p), and (ii) how much information departing from the node p (q) returns to that node p (q) without arriving at its destination. That is, the communication efficiency increases with the amount of information which departs from the originator and arrives at its destination, but decreases with the amount of information which is frustrated due to the fact that the information returns to its originator without being delivered to its target. This has lead to the definition (3.1).

It has been indeed proved that the function ξ_{pq} is a Euclidean distance between the nodes p and q in Γ [11].

THEOREM 3.1 [19]. *The communicability distance ξ_{pq} induces an embedding of the graph Γ of size n into a hypersphere of radius $R^2 = [c - (2 - b)^2/a]/4$ in an $(n - 1)$ -dimensional space, where $a = \vec{1}^T e^{-A} \vec{1}$, $b = \vec{s}^T e^{-A} \vec{1}$ and $c = \vec{s}^T e^{-A} \vec{s}$ with $\vec{s} = \text{diag } e^A$.*

Let us hereafter give a geometric view of the communicability distance. We first prove the following theorem.

THEOREM 3.2. *Let $\vec{x}_p = e^{\Lambda/2} \vec{\phi}_p$, where $\vec{\phi}_p = (\psi_1(p) \ \cdots \ \psi_\mu(p) \ \cdots \ \psi_n(p))^T$. Then we have*

$$(3.2) \quad G_{pq} = \vec{x}_p \cdot \vec{x}_q.$$

Proof. Let $X = (\vec{x}_1 \ \cdots \ \vec{x}_p \ \cdots \ \vec{x}_n) = e^{\Lambda/2} U^T$. We therefore have

$$(3.3) \quad X^T X = U e^{\Lambda} U^T = e^A = G,$$

which is immediately followed by Eq. (3.2). \square

This theorem transforms the communicability distance (3.1) into the form

$$(3.4) \quad \xi_{pq}^2 = \vec{x}_p \cdot \vec{x}_p + \vec{x}_q \cdot \vec{x}_q - 2\vec{x}_p \cdot \vec{x}_q = (\vec{x}_p - \vec{x}_q)^2.$$

In other words, the communicability distance is the Euclidean distance in the space of $\{\vec{x}_p\}$, which we will visualize hereafter.

For this purpose, we express \vec{x}_p in another way, using the identity $e^{\Lambda/2} = U^T e^{A/2} U$. We thereby have $X = U^T e^{A/2}$, which gives the μ th element of the vector \vec{x}_p as

$$(3.5) \quad (\vec{x}_p)_\mu = X_{\mu p} = \vec{\psi}_\mu \cdot (e^{A/2} \vec{w}_p),$$

where \vec{w}_p has 1 in the p th element and zero elsewhere.

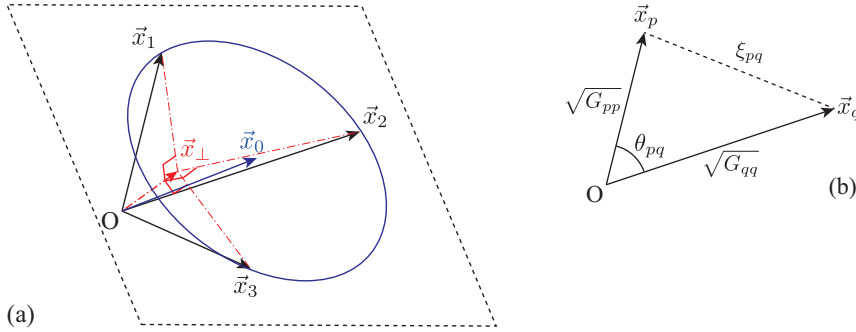


FIG. 1. (a) Three vectors \vec{x}_1 , \vec{x}_2 and \vec{x}_3 (solid black arrows) in a three-dimensional space spanned by the three eigenvectors of a 3×3 adjacency matrix A . The vectors fall on a two-dimensional flat surface (broken black lines) to which the vector \vec{x}_\perp (red dot-dashed arrow) is normal. We can draw a circle (solid blue curve) on the two-dimensional surface around a point \vec{x}_0 (solid blue arrow) to contain all three points. (b) The triangle spanned by the vectors \vec{x}_p and \vec{x}_q .

We can visualize this in the following way. Suppose that we place particles on the p th node and let them disperse according to the evolution operator $e^{A/2}$; we end up with a particle distribution all over the network. The vector $e^{A/2}\vec{w}_p$ specifies the particle distribution. We then break it down into the eigenmodes $\{\vec{\psi}_\mu\}$ in the form

$$(3.6) \quad e^{A/2}\vec{w}_p = \sum_{\mu=1}^n c_\mu \vec{\psi}_\mu.$$

The amplitude c_μ is given by $(\vec{x}_p)_\mu$. This tells us that the space of the vectors $\{\vec{x}_p\}$ is an n -dimensional one spanned by the eigenvectors $\{\vec{\psi}_\mu\}$.

Theorem 3.1 dictates that the vectors $\{\vec{x}_p\}$ fall onto the surface of a hypersphere in the space; see Fig. 1(a) for illustration in the case $n = 3$. We can understand this in the following way. We first fix the n -dimensional normal vector \vec{x}_\perp from n pieces of conditions $(\vec{x}_p - \vec{x}_\perp) \cdot \vec{x}_\perp = 0$ for $1 \leq p \leq n$. It specifies the $(n-1)$ -dimensional flat surface on which all vectors fall as $(\vec{x} - \vec{x}_\perp) \cdot \vec{x}_\perp = 0$. We next fix the n -dimensional vector \vec{x}_0 that specifies the center of the hypersphere as well as the radius R from $n+1$ pieces of conditions $(\vec{x}_0 - \vec{x}_\perp) \cdot \vec{x}_\perp = 0$ and $|\vec{x}_p - \vec{x}_0| = R$ for $1 \leq p \leq n$.

We can therefore regard ξ_{pq} as the cord distance between the two points on the hypersurface. Figure 1(b) picks out the triangle spanned by the vectors \vec{x}_p and \vec{x}_q . This leads to the definition in the next section of the angle between the two vectors.

4. Communicability angle. Let p and q be nodes of a connected simple network and let us define the following quantity:

$$(4.1) \quad \gamma_{pq} := \frac{G_{pq}}{\sqrt{G_{pp}G_{qq}}}.$$

We then prove the following result.

THEOREM 4.1. *The index γ_{pq} is the cosine of the Euclidean angle spanned by the position vectors of p and q .*

Proof. The view shown in Fig. 1(b) obviously gives

$$(4.2) \quad \cos \theta_{pq} = \frac{\vec{x}_p \cdot \vec{x}_q}{|\vec{x}_p| |\vec{x}_q|}.$$

The use of Eq. (3.2) then proves the result. \square

We then call θ_{pq} the communicability angle between the corresponding nodes of the graph. For each pair of nodes in the graph, the communicability distance and angle are related mathematically by the following expression:

$$(4.3) \quad \xi_{pq}^2 = G_{pp} + G_{qq} - 2\sqrt{G_{pp}G_{qq}} \cos \theta_{pq}$$

Because $G_{pq} \geq 0$ for any pair of nodes in Γ , the communicability angle is bounded by $0 \leq \cos \theta_{pq} \leq 1$. That is, the communicability angle of simple graphs can take values only in the range $(0^\circ, 90^\circ)$. We will now give classes of graphs that show how we attain the extremal values.

PROPOSITION 4.2. *Let P_n be the path graph with n nodes labeled by $1, 2, \dots, n$ sequentially. The communicability angle between any pair of nodes in P_n is given by*

$$(4.4) \quad \cos \theta_{pq}(P_n) = \frac{I_{p-q}(2) - I_{p+q}(2)}{\sqrt{[I_0(2) - I_{2r(p)}(2)][I_0(2) - I_{2r(q)}(2)]}}$$

in the limit $n \rightarrow \infty$, where $I_\gamma(z)$ is the Bessel function of the first kind and

$$(4.5) \quad r(p) = \begin{cases} p & \text{for } p \leq n/2 \text{ with even } n \text{ or } p \leq (n+1)/2 \text{ with odd } n, \\ n-p+1 & \text{for } p > n/2 \text{ with even } n \text{ or } p > (n+1)/2 \text{ with odd } n. \end{cases}$$

Proof. The eigenvalues and eigenvectors of the adjacency matrix of P_n are

$$(4.6) \quad \lambda_j(P_n) = 2 \cos \frac{j\pi}{n+1}, \quad \psi_j(p) = \sqrt{\frac{2}{n+1}} \sin \frac{jp\pi}{n+1}$$

for $1 \leq j \leq n$. Thus

$$(4.7) \quad G_{pq}(P_n) = \frac{1}{n+1} \sum_{j=1}^n \left[\cos \frac{j\pi(p-q)}{n+1} - \cos \frac{j\pi(p+q)}{n+1} \right] e^{2 \cos(j\pi/(n+1))},$$

$$(4.8) \quad G_{pp}(P_n) = \frac{1}{n+1} \sum_{j=1}^n \left[1 - \cos \frac{2j\pi p}{n+1} \right] e^{2 \cos(j\pi/(n+1))}.$$

In the limit $n \rightarrow \infty$, we can write them in integral forms, which eventually reduce to $G_{pq}(P_n) = I_{p-q}(2) - I_{p+q}(2)$, and $G_{pp}(P_n) = I_0(2) - I_{2r(p)}(2)$; it proves Eq. (4.4). \square

Notice that for the pair of nodes at the ends of the path we have

$$(4.9) \quad \lim_{n \rightarrow \infty} \cos \theta_{n1}(P_n) = \lim_{n \rightarrow \infty} \frac{I_{n-1}(2) - I_{n+1}(2)}{I_0(2) - I_2(2)} = 0,$$

which attains the lower bound of the communicability angle.

PROPOSITION 4.3. *Let $K_{1,n-1}$ be the star graph with n nodes. Let the node with degree $n-1$ labelled as 1. The communicability angle between any pair of nodes in $K_{1,n-1}$ is given by*

$$(4.10) \quad \cos \theta_{1q}(K_{1,n-1}) = \frac{\tanh^2(\sqrt{n-1})}{(n-2) \operatorname{sech}(\sqrt{n-1}) + 1} \quad \text{for } q \neq 1,$$

$$(4.11) \quad \cos \theta_{pq}(K_{1,n-1}) = \frac{\cosh(\sqrt{n-1}) - 1}{(n-2) \cosh(\sqrt{n-1}) + n-2} \quad \text{for } p \neq 1 \text{ and } q \neq 1.$$

Proof. The communicability between the different pairs of nodes in $K_{1,n-1}$ are

$$(4.12) \quad G_{1q}(K_{1,n-1}) = \frac{1}{\sqrt{n-1}} \sinh(\sqrt{n-1}) \quad \text{for } q \neq 1,$$

$$(4.13) \quad G_{pq}(K_{1,n-1}) = \frac{1}{n-1} [\cosh(\sqrt{n-1}) - 1] \quad \text{for } p \neq 1 \text{ and } q \neq 1.$$

The subgraph centrality of the two distinct nodes in the star graph are

$$(4.14) \quad G_{11}(K_{1,n-1}) = \cosh(\sqrt{n-1}),$$

$$(4.15) \quad G_{pp}(K_{1,n-1}) = \frac{1}{n-1} [\cosh(\sqrt{n-1}) + n - 2] \quad \text{for } p \neq 1.$$

Algebra with trigonometric identities gives Eqs. (4.10) and (4.11). \square

It is important to notice that

$$(4.16) \quad \lim_{n \rightarrow \infty} \cos \theta_{1q}(K_{1,n-1}) = 1 \quad \text{for } q \neq 1,$$

$$(4.17) \quad \lim_{n \rightarrow \infty} \cos \theta_{pq}(K_{1,n-1}) = 1 \quad \text{for } p \neq 1 \text{ and } q \neq 1,$$

which attain the upper bound of the communicability angle.

PROPOSITION 4.4. *Let K_n be the complete graph with n nodes. The communicability angle between any pair of nodes in K_n is given by*

$$(4.18) \quad \cos \theta_{pq} = \frac{e^n - 1}{e^n + n - 1}.$$

Proof. The eigenvalues of the adjacency matrix of K_n are $n - 1$ with multiplicity 1 and -1 with multiplicity $n - 1$. We thereby have

$$(4.19) \quad G_{pp} = \frac{1}{ne} (e^n + n - 1), \quad G_{pq} = \frac{1}{ne} (e^n - 1)$$

which proves Eq. (4.18). \square

Notice that $\cos \theta_{pq} \rightarrow 1$ as $n \rightarrow \infty$ in K_n .

5. Communicability distance and communicability angle. An interesting difference between the communicability distance ξ_{pq} and the communicability angle θ_{pq} arises from their analysis in a path P_n . First, we prove the following result for the communicability distance.

PROPOSITION 5.1. *Let P_n be a path graph of n nodes labeled consecutively from one end point to the other as $1, 2, \dots, n$. Let $S = \{\xi_{12}^2, \xi_{13}^2, \dots, \xi_{1n}^2\}$ be the ordered sequence of communicability distances between the first node and any other nodes q in the path. Then, S is nonmonotonic.*

Proof. Without any loss of generality we will consider here even n for simplicity. The communicability distance in question is given by

$$(5.1) \quad \xi_{1q}^2 = \begin{cases} [2I_0(2) - I_2(2)] - [I_{2q}(2) + 2I_{1-q}(2) - 2I_{1+q}(2)] & \text{for } 1 < q \leq n/2, \\ [2I_0(2) - I_2(2)] - [I_{2(n-q+1)}(2) + 2I_{1-q}(2) - 2I_{1+q}(2)] & \text{for } q > n/2, \end{cases}$$

where $r(p)$ and $I_\gamma(z)$ are as before. First, we have

$$(5.2) \quad \xi_{12}^2 \simeq 1.0637$$

in the limit $n \rightarrow \infty$. Next, let $\chi(q) = I_{2q}(2) + 2I_{1-q}(2) - 2I_{1+q}(2)$. It is easy to check that $\chi(q) > \chi(q+1)$, so that ξ_{1q}^2 increases as $q \rightarrow n/2$. For nodes relatively close to the center of the path, we have

$$(5.3) \quad \lim_{q \rightarrow n/2} \xi_{1q}^2 = 2I_0(2) - I_2(2) \approx 3.8702,$$

but as q approaches the other end of the path, we have

$$(5.4) \quad \lim_{q \rightarrow n} \xi_{1q}^2 = 2I_0(2) - 2I_2(2) \approx 3.1813.$$

This means that the communicability distances increases from ξ_{12} up to the maximum $\xi_{1q} \approx 3.8702$ and then decreases to $\xi_{1n} \approx 3.1813$, which proves the result. \square

We now prove that the monotonicity holds for the communicability angle.

PROPOSITION 5.2. *Let P_n be a path graph of n nodes labeled consecutively from one end point to the other as $1, 2, \dots, n$. Let $C = \{\theta_{12}, \theta_{13}, \dots, \theta_{1n}\}$ be the ordered sequence of communicability angles between the first node and any other nodes q in the path. Then, C is monotonic.*

Proof. Without any loss of generality we will consider here again even n . The communicability angle in question is given by:

$$(5.5) \quad \cos \theta_{1q} = \begin{cases} \frac{I_{1-q}(2) - I_{1+q}(2)}{\sqrt{[I_0(2) + I_2(2)][I_0(2) - I_{2q}(2)]}} & \text{for } 1 < q \leq n/2, \\ \frac{I_{1-q}(2) - I_{1+q}(2)}{\sqrt{[I_0(2) + I_2(2)][I_0(2) - I_{2(n-q+1)}(2)]}} & \text{for } q > n/2. \end{cases}$$

For small values of q it is easy to see that $\cos \theta_{1q} > \cos \theta_{1,q+1}$; the numerator of (5.5) decreases as q increases and at the same time the denominator decreases. It is also easy to see that $\lim_{q \rightarrow \infty} \cos \theta_{1q} = 0$.

The difference with the result for the communicability distance arises from the fact that the numerator of (5.5) for $q > n/2$ is the same as that for $1 < q \leq n/2$. We therefore have $\lim_{q \rightarrow \infty} \cos \theta_{1q} = 0$ for $q > n/2$, which indicates that once the angle between the first and the q th nodes in P_n reaches its maximum value, *i.e.*, 90° , it does not decrease again, which proves that the series C is monotonic. \square

Now, let us extract the structural information provided by these results which will be useful for further application of the communicability angle in analyzing real-world complex networks. Let us define the average communicability angle for a given graph as the average over the pairs of nodes:

$$(5.6) \quad \langle \theta \rangle = \frac{2}{n(n-1)} \sum_{p>q} \theta_{pq}.$$

We then have the following observations: (i) The average communicability angle for the path graph P_n tends to 90° when the number of nodes tends to infinite. This is a consequence of Propositions 4.2 and 5.2; (ii) The average communicability angle for the star graph $K_{1,n-1}$ tends to 0° when the number of nodes tends to infinite. This is a consequence of Proposition 4.3; (iii) The average communicability angle for the complete graph K_n tends to 0° when the number of nodes tends to infinite. This is a consequence of Proposition 4.4.

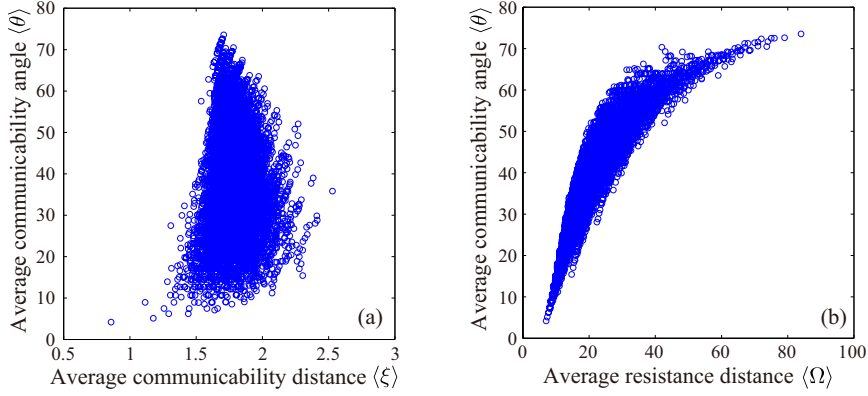


FIG. 2. Scatter plot of the average communicability angle against (a) the average communicability distance and (b) the average resistance distance for all 11,117 connected graphs with 8 nodes.

6. Computational analysis of the communicability angle. In this section we computationally analyze the average communicability angle (5.6) for connected graphs. Specifically, we here study a dataset of all 11,117 connected graphs with 8 nodes. We divide this section into three subsections: we first analyze relations (or lack thereof) between the average communicability angle and other graphs metrics, namely the average path length, the average resistance distance and the average communicability distance; we then study relations between $\langle \theta \rangle$ and the graph planarity; we finally investigate influence of graph modularity on the communicability angle.

6.1. Communicability angle and other graph metrics. We first compare the average communicability angle $\langle \theta \rangle$ with the average communicability distance $\langle \xi \rangle$, the average path length $\langle l \rangle$ and the average resistance distance $\langle \Omega \rangle$ as metrics potentially related to $\langle \theta \rangle$; every average was taken over all pairs of nodes. We show in Fig. 2 the comparison with the average communicability and resistance distances. The results for the average path length is very similar and not shown here.

We can see that the communicability angle is not directly or trivially related to the other metrics. It is particularly interesting to see the lack of correlation between $\langle \theta \rangle$ and $\langle \xi \rangle$. The average communicability angle shows more similar trends to the average path length $\langle l \rangle$ and the average resistance distance $\langle \Omega \rangle$. The extreme values of $\langle l \rangle$ and $\langle \Omega \rangle$ coincide with those of $\langle \theta \rangle$, although there is a large dispersion in between.

The main conclusion of this subsection is that the communicability angle is not trivially dependent on other graph metrics. This is potentially important in applications because it indicates that the average communicability angle accounts for a new kind of structural information of graphs which is not accounted for by the other metrics. Among all the connected graphs with 8 nodes, the path graph P_8 has the largest average communicability angle and the complete graph K_8 has the smallest. Among all the trees with 8 nodes, the star graph $K_{1,7}$ has the smallest average communicability angle. This is also verified for all connected graphs with 5, 6 and 7 nodes. We thereby have the following:

CONJECTURE 6.1. *Among all connected graphs with n nodes, the average communicability angle is the largest for the path graph P_n and the smallest for the complete graph K_n .*

CONJECTURE 6.2. *Among all trees with n nodes, the average communicability angle is the largest for the path graph P_n and the smallest for the star graph $K_{1,n-1}$.*

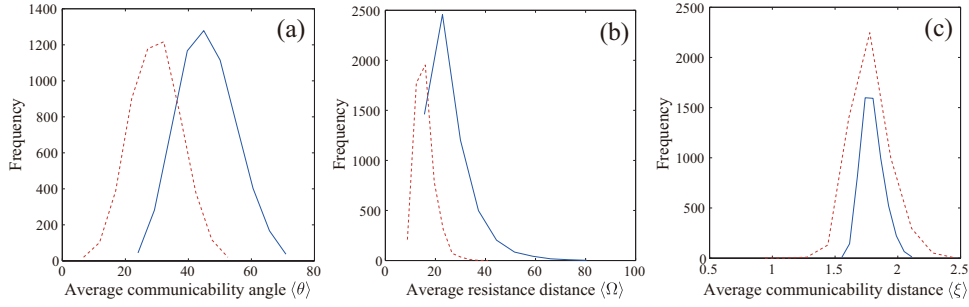


FIG. 3. Probability of finding planar graphs among the graphs having certain values of (a) the average communicability angle, (b) the average resistance distance, and (c) the average communicability distance. The histogram for planar graphs is displayed as a solid line and that for nonplanar graphs as a broken line.

These observations indicate that the average communicability angle describes the efficiency of a graph in using the space in which it is embedded. The path graph P_n , which intuitively occupies the largest portion of space, has the largest average communicability angle, while the star and complete graphs, which intuitively occupy the smallest, have the average communicability angle close to zero. In the next section we explore more observations of this sort from a computational point of view.

6.2. Communicability angle and graph planarity. Here we investigate the relation between the graph planarity and the average communicability angle. We first determine whether a graph is planar or not using the algorithm of Boyer-Myrvold planarity test [5]. We then construct the histogram of the frequency of planar/nonplanar graphs with respect to the average communicability angle.

Let η_k be the number of planar graphs having $k \leq \langle\theta\rangle < (k + 10^\circ)$ for $k = 0^\circ, 10^\circ, 20^\circ, \dots, 80^\circ$. We plot in Fig. 3 the histogram of the planar/nonplanar graphs as a function of their values of $\langle\theta\rangle$ for all connected graphs with 8 nodes. For comparison, we also show similar plots for the average resistance and communicability distances. The plot for the average path length is very similar to the one for the resistance distance and is not shown.

The first interesting observation is that the planar graphs display significantly smaller values of $\langle\theta\rangle$ than the nonplanar graphs. The peaks in the histogram Fig. 3(a) for the planar and nonplanar graphs are at $\langle\theta\rangle \approx 29.5$ and $\langle\theta\rangle \approx 45.8$, respectively. There is also a significant separation between the maxima of the two histograms, which contrasts very much with what is observed for the other metrics. This result does not necessarily mean that the average communicability angle characterizes the graph planarity or vice versa, but that the planarity is indeed an important ingredient of the communication efficiency as measured by the communicability angle.

The following are observations about the average communicability angle obtained from the analysis of the connected graphs with 8 nodes: (i) No planar graph has $\langle\theta\rangle < 21.4$; (ii) The planar graphs with the smallest value of $\langle\theta\rangle$ are derived from graphs having a subgraph homeomorphic to K_5 in which a link of the 5-nodes clique is deleted. Examples are given in Fig. 4; (iii) There is no nonplanar graph with $\langle\theta\rangle > 55.065$; (iv) The nonplanar graphs with the largest values of $\langle\theta\rangle$ are graphs which contains a subgraph homeomorphic to $K_{3,3}$. Examples are given in Fig. 5.

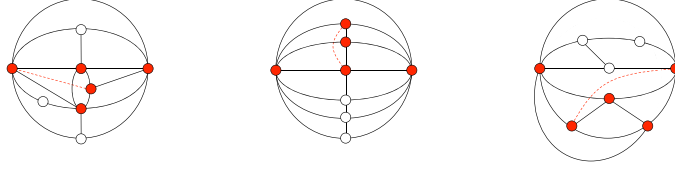


FIG. 4. Three planar graphs with 8 nodes which have the smallest values of $\langle \theta \rangle$. The graphs are drawn to emphasize their planarity, with the edges of the graphs represented by continuous black lines. The five nodes marked in red form a 5-clique if the edge marked as a dotted red line is added to the graph. The corresponding graph having the red dotted line added as an edge is nonplanar as there is a subgraph homeomorphic to K_5 .

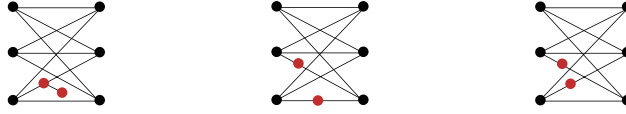


FIG. 5. Three nonplanar graphs with 8 nodes which have the largest values of $\langle \theta \rangle$. The graphs are drawn to emphasize the existence of subgraphs homeomorphic to $K_{3,3}$. The subgraph $K_{3,3}$ is drawn with the nodes in black.

6.3. Communicability angle and graph modularity. Modularity is a very important concept for the study of real-world networks. It refers to the property of graphs with clusters of highly interconnected nodes but with poor inter-cluster connectivity. Such clusters are usually referred to as communities in network theory and are expected to play fundamental organizational roles in real-world networks, *e.g.*, groups of proteins with similar actions and groups of people with common interests. A network with such clusters has structural bottlenecks; that is, if small groups of nodes/edges are removed the network is disconnected into two or more relatively large connected components. An extreme case are the dumbbell graphs K_n-K_n , that is, two cliques of n nodes connected by only one edge; the removal of the edge separates the network into two connected components of $n/2$ nodes each.

On the other hand, a graph without such bottlenecks shows a super-homogeneous structure, which is usually referred to as a good expansion property. Intuitively, an expander is characterized by the fact that every subset S with more than $n/2$ nodes has a large boundary, which is the number of edges with one node inside the set S and the other in \bar{S} [31]. Expander graphs are characterized by having a large spectral gap $\lambda_1 - \lambda_2$ of the adjacency matrix [1]; see Refs. [25, 28] for details.

What is important for the present subsection is that expanders are characterized by the lack of modularity, *i.e.*, the lack of tightly connected clusters which are poorly interconnected by structural bottlenecks. In networks where $\lambda_1 \gg \lambda_2$, we have the following expression for the communicability angle:

$$(6.1) \quad \cos \theta_{pq} = \frac{G_{pq}}{\sqrt{G_{pp}G_{qq}}} \simeq \frac{\psi_1(p) \psi_1(q) e^{\lambda_1}}{\sqrt{\psi_1(p)^2 e^{\lambda_1} \psi_1(q)^2 e^{\lambda_1}}} = \cos 0^\circ.$$

That is, the networks lacking any modularity are characterized by very small value of the communicability angle. On the other hand, in a network where λ_1 is not

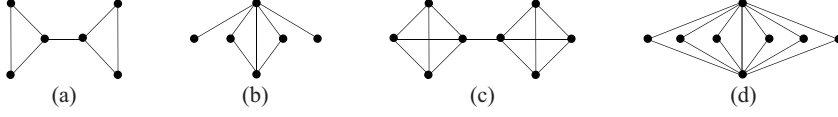


FIG. 6. The graphs with 6 nodes and 7 edges (a) with the largest and (b) the smallest average communicability angles. The same for the graphs with 8 nodes and 13 edges (c) and (d).

significantly larger than λ_2 , we make use of the expansions

$$(6.2) \quad G_{pp}G_{qq} = \psi_1(p)^2\psi_1(q)^2e^{2\lambda_1} + (\psi_1(p)^2\psi_2(q)^2 + \psi_2(p)^2\psi_1(q)^2)e^{\lambda_1+\lambda_2} \\ + \psi_2(p)^2\psi_2(q)^2e^{2\lambda_2} + \text{h.o.},$$

$$(6.3) \quad G_{pq}^2 \simeq \psi_1(p)^2\psi_1(q)^2e^{2\lambda_1} + 2\psi_1(p)\psi_1(q)\psi_2(p)\psi_2(q)e^{\lambda_1+\lambda_2} \\ + \psi_2(p)^2\psi_2(q)^2e^{2\lambda_2} + \text{h.o.},$$

where h.o. denotes the higher-order terms. The communicability angle is thereby transformed into the form

$$(6.4) \quad \cos \theta_{pq} = \frac{G_{pq}}{\sqrt{G_{pq}^2 + (\psi_1(p)\psi_2(q) - \psi_2(p)\psi_1(q))^2 e^{\lambda_1+\lambda_2} + \text{h.o.}}}.$$

The second term in the denominator depends on the size of the spectral gap; the closer λ_2 is to λ_1 , *i.e.*, the smaller the spectral gap, the larger the denominator is, and consequently, the smaller Eq. (6.4) is. Therefore the angle θ_{pq} is larger as the spectral gap is smaller. We should remark here that θ_{pq} does not depend only on the spectral gap because the higher-order terms in Eq. (6.4) can make an important contribution.

Let us show examples that illustrate the above important relation between the communicability angle and the graph modularity. Here again we focus on $\langle \theta \rangle$. We first consider the dumbbell graph K_3 - K_3 shown in Fig. 6(a). It consists of two cliques of 3 nodes each, which are connected by a link, thus having 7 edges in total. The average communicability angle for this graph is $\langle \theta \rangle \approx 57.105$ and its spectral gap is $\Delta \approx 0.682$. Among the 19 graphs with 6 nodes and 7 edges, the dumbbell K_3 - K_3 has the largest value of $\langle \theta \rangle$. The smallest value of the average communicability angle is obtained for the graph in Fig. 6(b), having $\langle \theta \rangle \approx 47.935$ and $\Delta \approx 2.284$.

The situation is very similar for the 1,454 graphs with 8 nodes and 13 edges, among which the dumbbell graph K_4 - K_4 in Fig. 6(c) has the largest average communicability angle $\langle \theta \rangle \approx 53.876$ with the spectral gap $\Delta \approx 0.511$. The graph with the smallest value of $\langle \theta \rangle$ is the so-called agave graph shown in Fig. 6(d); it consists of two connected nodes each of which is also connected to the other $n-2$ nodes that are not connected among them. It has $\Delta = 4.00$ and $\langle \theta \rangle \approx 31.782$. The graphs with the second and third smallest average communicability angles, $\langle \theta \rangle \approx 35.123$ and $\langle \theta \rangle \approx 35.606$ with $\Delta \approx 2.988$ and $\Delta \approx 3.337$, respectively, have structures similar to the agave graph. Notice that the agave graph can be disconnected by removing two edges, but the remaining principal connected component has $n-1$ nodes, while the removal of 50% of the edges in this graph creates a principal connected component still containing 62.5% of the nodes. This shows the robustness of this graph to edge removal, a characteristic of good expansion graphs due to the lack of structural bottleneck.

Figure 7(a-b) shows planar embeddings of the graphs in Fig. 6(a-b), respectively, onto triangular lattices. The shadowed areas indicate the triangles covered by the

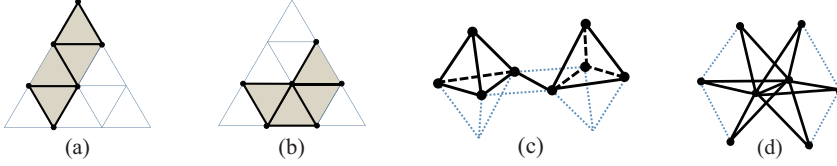


FIG. 7. (a–b) Planar embeddings of the graphs in Fig. 6(a–b), respectively, onto triangular lattices. (c–d) Three-dimensional embeddings of the graphs in Fig. 6(c–d), respectively, onto close-packed lattices.

graphs in these embeddings. Although both cover the four triangles, the latter graph, the one with the smallest average communicability angle, covers the most efficient packing in two-dimensional space, which is the area with a node surrounded by six others forming a hexagon. This is known as the penny-packing problem; see Ref. [22] for further information. The embedding of the graph with higher modularity and the largest average communicability angle is far from this optimal configuration.

A similar situation occurs with the graphs in Fig. 6(c–d), the ones with the largest and smallest $\langle\theta\rangle$ among those with 8 nodes and 13 edges; Fig. 7(c–d) show their embeddings onto close-packed lattices. We can conclude from these observations that a large average communicability angle indicates a poor spatial efficiency of the graph, while a small value of $\langle\theta\rangle$ is associated to the efficient use of space.

6.4. Conclusions of the computational analysis of simple graphs. The main conclusion of Section 6 is that the average communicability angle describes very well a graph characteristics which is related to their spatial efficiency. It is drawn from the following observations. First, planar graphs are not spatially efficient graphs; at the same time they have large average communicability angles. On the contrary, highly nonplanar graphs more efficiently use the available space; at the same time they have smaller values of $\langle\theta\rangle$. Second, a modular graph uses the available space less effectively than a nonmodular one; at the same time, modular graphs have relatively large values of the average communicability angle.

We should, however, be careful in analyzing more complex situations in which combinations of properties, such as planarity and modularity, are present. In general, we consider that graphs with relatively small values of the average communicability angle exhibit higher spatial efficiency than those with relatively larger values.

7. Communicability angle in real-world networks. We start this section by considering the average communicability angle of a series of 120 complex networks arising from various scenarios. The networks are briefly described in Supplementary Information accompanying this paper, where references to the original works are provided. The series includes networks in which the nodes and links are clearly embedded into geometrical spaces, such as urban street networks, networks formed by animal nests, brain and neural networks, protein-residue networks as well as electronic circuits and the Internet. It also includes networks in which the nodes and links can hardly be allocated to geographic positions, such as food webs, social networks and software networks. The biomolecular networks including protein-protein interaction and gene transcription networks are also non-geographically embedded ones.

7.1. Global properties of the communicability angle. The 120 real-world networks studied here cover the whole spectrum of values of the average communicability angle from $\langle\theta\rangle \approx 10^{-5}$ for the food web of Shelf to $\langle\theta\rangle \approx 89.9$ for the Power

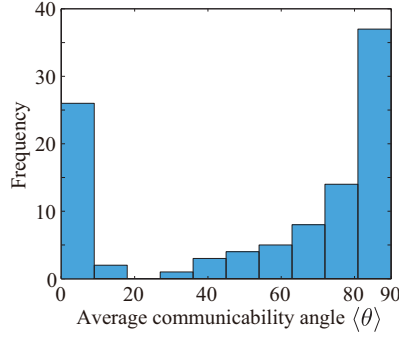


FIG. 8. Histograms of the average communicability angle in 120 real-world networks with the bin size of 9° .

Grid network of western USA. The histogram in Fig. 8 shows two prominent peaks at $0 \leq \langle \theta \rangle \leq 9$ and at $81 \leq \langle \theta \rangle \leq 90$. A more detailed view (not shown) indicates that the highest frequency occurs at $0 \leq \langle \theta \rangle \leq 1$, followed by the one at $89 \leq \langle \theta \rangle \leq 90$. That is, the real-world networks are very much polarized into the two extremes; either they have very small values of the communicability angle or very large ones.

Certain classes of networks have a large homogeneity in the values of the average communicability angle. The 1997 and 1998 versions of the Internet at Autonomous System (AS) have the average communicability angles of 0.78 and 0.42, respectively. There is also a large homogeneity among the brain/neural networks, namely, the visual-cortex networks of cat and macaque as well as the neural network of *C. elegans*, which have $\langle \langle \theta \rangle \rangle = 1.77 \pm 1.66$, where the brackets $\langle \langle \cdot \cdot \cdot \rangle \rangle$ denote the average value of the average communicability angles for a series of networks. In addition, the classes of urban street networks formed by 14 networks and the one of protein-residue networks formed by 40 networks also show remarkable homogeneity. For instance, the urban street networks have $\langle \langle \theta \rangle \rangle = 86.07 \pm 5.07$ and the protein-residue networks have $\langle \langle \theta \rangle \rangle = 78.83 \pm 7.28$. The ranking of the 14 cities in the former is: Barcelona < Rio Grande < Yuliang < Chegkan < Atlanta < Berlin < Rotterdam < Hong Kong < Mecca < Cambridge < Oxford < Penang < Ahmedabad < Milton Keynes. This means that in terms of the effective communication among the different regions of the city, Barcelona is the most effective one, while Milton Keynes the worse.

The homogeneity among the protein-residue networks is more unexpected than that among the urban street networks because they represent three-dimensional (3D) objects. Proteins are folded into 3D structures forming topologies consisting of mainly α -helices, mainly β -sheets, or both mixed. They also have different shapes and sphericities. It is therefore surprising that the protein-residue networks are characterized by very large values of the communicability angle, which are more characteristic of planar or almost planar networks, as demonstrated for the urban street networks.

Although we will go back below to the relation between the communicability angle and the structure of proteins, let us make a comment here. The fact that proteins are embedded into the 3D physical space does not necessarily mean that their residue networks are nonplanar. The same applies to other naturally evolving networks, such as the networks of galleries and corridors formed by termite mounds, which are also characterized by very large average communicability angles with $\langle \langle \theta \rangle \rangle = 88.33 \pm 1.01$. Although the mounds are constructed in the 3D space, they are remarkably close to planar graphs; we have indeed found that by removing only 6% of the edges of these

networks the graphs representing them become planar. Both the termite mounds and the protein-residue networks have certainly evolved in the 3D space, but the networks must be close to planar graphs for different ecological or biological reasons. In the termite mounds the use of a large volume of the 3D space is needed to produce a ventilation system necessary to discharge the carbon dioxide produced in its interior. For protein, structures close to planar ones are needed to avoid high compactness that destroy the internal cavities of the protein needed for developing their functions; see Section 7.2 below.

On the other hand, the values of $\langle\theta\rangle$ obtained for the software networks [29] are unexpectedly heterogeneous. These networks yield $\langle\langle\theta\rangle\rangle = 57.6 \pm 30.7$ with the values ranging from $\langle\theta\rangle \approx 3.465$ for Linux to $\langle\theta\rangle \approx 84.323$ for XMMS. The ranking of these networks in terms of the average communicability angle is: Linux < MySQL < VTK < Abi Word < Digital Material < XMMS. The classes of social and biological networks consisting of 14 and 11 networks, respectively, also show relatively large variability in their values of the communicability angle: $\langle\langle\theta\rangle\rangle = 55.8 \pm 21.3$, and $\langle\langle\theta\rangle\rangle = 63.3 \pm 17.0$, respectively. This is not surprising; we can easily associate it to the diversity of networks in these classes.

What is really surprising is that the food webs, which form a very homogeneous class of networks in terms of the relations accounted for them, yield a relatively large standard deviation in the values of the communicability angle: $\langle\langle\theta\rangle\rangle = 7.1 \pm 16.1$ with the values ranging from $\langle\theta\rangle \approx 10^{-5}$ for the marine system of Shelf to $\langle\theta\rangle \approx 78.356$ for the web of the English grassland. The ranking of these food webs in terms of the average communicability angle is: Shelf < Elverde < Skipwith < ReefSmall < LittleRock < Stony < Coachella < Canton < Benguela < BridgeBrook < Ythan2 < Ythan1 < StMartins < StMarks < ScotchBroom < Chesapeake < Grassland.

In terms of the individual values of $\langle\theta\rangle$, the results obtained for these 120 networks agree with our findings in the previous section. The largest average communicability angles are observed for the Power Grid of western USA and urban street networks, which are planar or almost planar with both nodes and edges embedded into a plane. On the other extreme of the smallest average communicability angles, there are networks which are highly nonplanar, such as the USA air transportation network, a world trade network, the Internet at AS, and brain/neural networks. All these networks have nodes embedded into two- or three-dimensional spaces, such as cities, countries or organs, but the edges connecting them very efficiently use the available space. We would like to remark here that the small values of $\langle\theta\rangle$ observed in some classes of networks do not necessarily mean a high interconnection density. For instance, the USA airport transportation network and the two versions of the Internet studied here have relatively small edge densities: 0.039 and 0.0011, respectively.

7.2. Communicability angle and spatial efficiency of proteins. We have accumulated several pieces of empirical evidence that support the idea that the average communicability angle accounts for the spatial efficiency of graphs. It is, however, generally difficult to find quantitative measures of the spatial efficiency in real-world complex networks to compare with the communicability angle.

An exception to this is provided by proteins, which are 3D objects characterized by different degrees of packing or spatial efficiency. In this section we study the relation between the average communicability angle and the spatial efficiency of the protein-residue networks for a group of 40 proteins whose 3D structures have been resolved by X-ray crystallography and deposited in the protein data bank (PDB) [4]. Here each node represents an amino acid in the protein and two nodes are connected

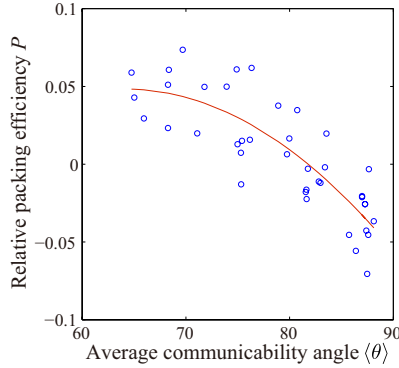


FIG. 9. Linear correlation between the average communicability angle of proteins represented by residue networks and the relative packing efficiency.

if the corresponding amino acids are separated at a distance of no more than 7\AA in the 3D structure of the protein as determined experimentally [2].

A protein is a linear sequence of amino acids connected by peptide bonds. The chain is folded into a 3D shape unique to each protein. While the amino-acid sequence forms the so-called primary structure of the protein, the 3D folding defines its secondary and tertiary structures. The secondary one is characterized by the presence of the α -helices and the β -sheets, while the tertiary one is formed by global positioning of the secondary one into a 3D shape that gives the protein its globular-like structure [9]. The folding of the proteins is the consequence, *grosso modo*, of two main necessities that the protein has: (i) protecting the hydrophobic amino acids from their contact with water; (ii) occupying a minimum space inside the limited volume of the cell. Thus the packing of a protein is related to its spatial efficiency [20], which is responsible for many of its physico-chemical and biological properties.

There are many ways of quantifying the packing of a protein, but here we consider the following one. Let V_e be the volume of a protein which is expected from its ideal 3D structure and let V_o be the volume which is actually observed in its X-ray crystallography. We then define the relative deviation from its ideal volume as

$$(7.1) \quad P = \frac{V_e - V_o}{V_e}.$$

Hereafter we call P the relative packing efficiency of the protein. A positive value of P means that the protein is more packed than expected from its ideal 3D structure, that it is highly efficient in using the 3D space, at least relatively to the ideal structure. A negative value of P , on the other hand, means that it is less packed than expected, that it is not spatially efficient. We should mention here that values that deviate very much from the expected or ideal values can indicate possible problems with the structure and as such should be discarded from the analysis.

Using computational techniques and VADAR software described in Ref. [33], we have calculated the expected and observed volumes of the 40 proteins. We show in Fig. 9 the relation between the relative packing efficiency P and the average communicability angle of the 40 proteins. The Pearson correlation coefficient is $R = -0.837$, indicating a significant correlation between the two variables. We can summarize the results as follows: (i) proteins with poor spatial efficiency, $P < 0$, have $\langle \theta \rangle > 81^\circ$; (ii) those with high spatial efficiency, $P > 0$, have $\langle \theta \rangle < 80^\circ$. In other words, small aver-

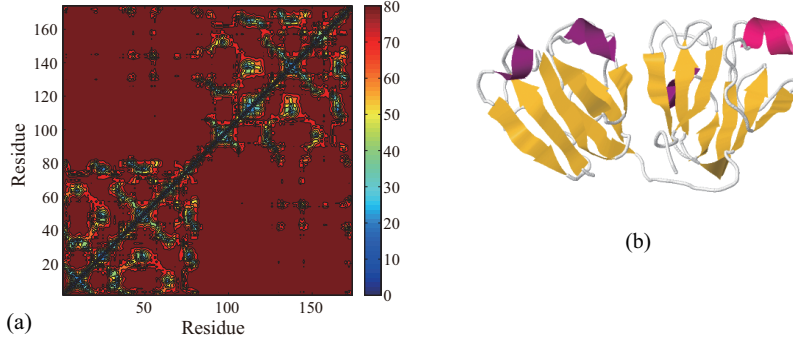


FIG. 10. (a) Contour plot of the communicability angle between every pair of residues in the GammaB crystallin protein with PDB code 1amm. (b) A cartoon representation of the protein with PDB code 1amm in which the β -sheets are represented as arrows in yellow and the helices as ribbons in magenta. This shows the existence of two domains in it.

age communicability angles are related to high spatial efficiency of proteins while large average communicability angles with a poor use of space. We notice in passing that there are no proteins with $\langle \theta \rangle < 60^\circ$, which can be explained by the fact that increasing too much packing would make the internal cavities of the protein disappear [20]. The internal cavities are indeed responsible for the interaction of proteins with other biological molecules and usually play a fundamental role in their functionality.

Possibilities which the communicability angle brings to the analyses of the structure of spatially embedded complex networks obviously go beyond the use of $\langle \theta \rangle$. For instance, the contour plot of the communicability angle for every pair of residues in a protein can reveal important properties of its 3D structure. Figure 10 shows an example of the protein with PDB code 1amm, which corresponds to the GammaB crystallin, whose crystallographic analysis was carried out at 150K. This protein consists of two α, β -domains, the first of which formed by amino acids 1-83 and the second by amino acids 84-174. The two domains are very well reflected in the contour plot Fig. 10(a) as two main diagonal blocks of relatively small communicability angles, which indicates good internal communication in each domain.

7.3. Spatial efficiency in networks under external stress. The communicability function has been previously generalized to consider an external stress to which the network is submitted. This external stress is accounted for by means of the so-called inverse temperature $\beta \equiv (k_B T)^{-1}$, where k_B is a constant and T is the temperature [14]. This analogy results from regarding that the whole network is submerged into a thermal bath of the inverse temperature β ; see [16, 10] for details. After equilibration in the bath, all edges of the network acquire a weight equal to β .

It is clear that when $\beta \rightarrow 0$, *i.e.*, as the temperature tends to infinite, the network becomes disconnected and there is no communication among any pair of nodes. This resembles a gas in which every node is an independent particle. On the other hand, when $\beta \rightarrow \infty$, *i.e.*, the temperature tends to zero, the weights of every edge becomes extremely large, which definitively increases the communication capacity among the pairs of connected nodes. The temperature thus plays a role of an empirical parameter which is useful in simulating effects of external stresses to which the network is submitted, such as different levels of social agitation, economical situations, environmental stress, variable physiological conditions, *etc.* Under this analogy, we generalized the

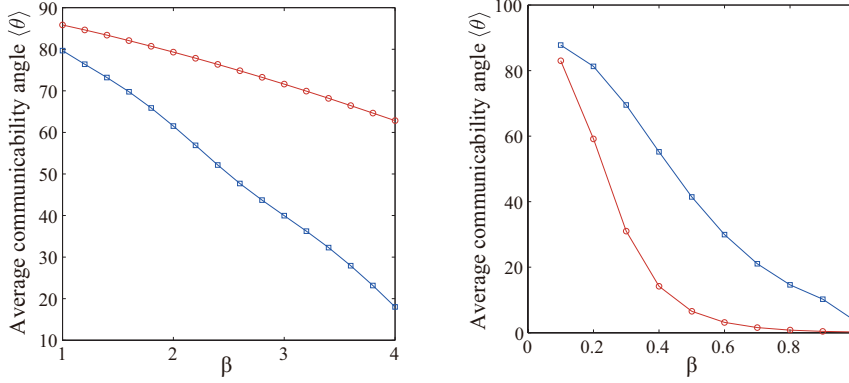


FIG. 11. (a) Effects of the inverse temperature β on the average communicability angle in two urban street networks, Rio Grande, Brazil (squares) and Yuliang, China (circles). (b) The same for two visual-cortex networks: cat (circles), macaque (squares).

communicability function (2.2) into the form [14]

$$(7.2) \quad G_{pq}(\beta) = (e^{\beta A})_{pq}.$$

It is straightforward to realize that the communicability angle between a given pair of nodes is generalized to

$$(7.3) \quad \cos \theta_{pq}(\beta) = \frac{G_{pq}(\beta)}{\sqrt{G_{pp}(\beta) G_{qq}(\beta)}}.$$

Let us conduct a simple experiment to explore the possibilities which this empirical parameter brings to the analysis of real-world scenarios. We use two urban street networks representing the city landscapes of Rio Grande in Brazil and of Yuliang in China. Both cities have large values of the average communicability angle, *i.e.*, small spatial efficiency, with $\langle \theta \rangle \approx 79.7$ and $\langle \theta \rangle \approx 85.8$, respectively. We then lower the temperature and see if it increases the spatial efficiency of both cities, *i.e.*, if it decreases the values of $\langle \theta \rangle$. In other words, we systematically increase β and compute the average communicability angle $\langle \theta(\beta) \rangle$. The increase in β here can be associated to the average increment in the number of lanes per street in the city.

Figure 11(a) shows the results. The city of Rio Grande dramatically improves its spatial efficiency by increasing the average number of lanes of its streets. Although the improvement for Yuliang is not so dramatic, there is still a decrease in the average communicability angle of 20° . The causes for the difference in the variation of $\langle \theta \rangle$ with the temperature for different networks is not a trivial one, as there should be many structural factors involved. We do not investigate these causes here.

We next carry out the opposite experiment using two brain networks representing the cat and macaque visual cortices. The average communicability angle shows that both networks have a great spatial efficiency: $\langle \theta \rangle \approx 0.22$ and $\langle \theta \rangle \approx 3.52$, respectively. We here raise the temperature, *i.e.*, decrease β , and see if it deteriorates the connections in the visual cortices in terms of the average communicability angle $\langle \theta(\beta) \rangle$. The decrease of β can be regarded as any malfunctioning or diseases.

Figure 11(b) shows the results. Both networks dramatically decrease their spatial efficiency as $\beta \rightarrow 0$; obviously, $\cos \theta_{pq}(\beta = 0) = 90^\circ$. We notice, however, that the

cat visual cortex is more resistant to the stress than the macaque one. For $\beta = 0.6$, for example, the former has $\langle\theta\rangle \approx 3.15$ while the latter has jumped up to $\langle\theta\rangle \approx 29.9$.

In closing, the use of the empirical parameter β allows us to simulate the effects of external factors which can modify the spatial efficiency of a network. This brings a modeling scenario to assaying of strategies of improving the spatial efficiency of networks or to analyses of their resilience to external stresses.

8. Conclusions. It would be argued that networks exist because of the necessity of communication among the entities of a complex system. Thus, communication is a main driver of the structural organization of complex networks. We have introduced here the concept of spatial efficiency of a network inspired by the similar one used in economics, where the spatial efficiency is used “*to characterize the ease with which economic activities are geographically organized and transacted within a region*” [32]. It particularly refers to “*the organization of physical assets, such as buildings, infrastructure, and green space, which structure the transportation, communication, public service, and energy needs of businesses and residents within the region and beyond*” [32].

In a network the more abstract spatial efficiency refers to the average quality of communication among the nodes. Such communication goodness is quantified as the ratio of the amount of information successfully delivered to its destination to the one which is frustrated in its delivery and returned to their originators. This new paradigm is then mathematically formulated in terms of the communicability angle between a pair of nodes. We have provided analytical and empirical pieces of evidence which reaffirm the idea that the communicability angle accounts for the spatial efficiency of networks.

The richness of this approach goes beyond the results presented here; there are a few immediate directions of research in this area which can open new opportunities for the analysis of networks. The use of the communicability angle for a pair of connected nodes can be seen as an edge centrality measure which may reveal important characteristics of individual edges in networks. The communicability angle averaged over the edges incident to a given node can also represent a node centrality index which indicates the contribution of the node to the global spatial efficiency of a network. The study of the effects of the inverse temperature on the spatial efficiency and the determination of the most important structural factors that influence it is of tremendous practical importance. These studies will allow us not only to predict the effects of external stresses over the spatial efficiency of a network but also to assay theoretical scenarios of improving this efficiency in certain classes of networks. Last but not least, the new concept of communicability angle can bring new possibilities to the mathematical analysis of specific types of graphs and properties, such as planarity and graph thickness among others.

Acknowledgement. EE thanks the Royal Society for a Wolfson Research Merit Award. He also thanks Dr. Sean Hanna (UCL) for the datasets of urban street networks used in this work.

REFERENCES

- [1] N. Alon and V. D. Milman, λ_1 , *isoperimetric inequalities for graphs, and superconcentrators*, J. Combin. Theor. B, 38 (1985) pp. 73–88.
- [2] A. R. Atilgan, P. Akan, and C. Baysal, *Small-world communication of residues and significance for protein dynamics*, Biophys. J. 86 (2004) pp. 85–91.

- [3] M. Barthélemy, *Spatial networks*, Phys. Rep., 499 (2011) pp. 1–101.
- [4] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, *The protein data bank*, Nucleic Acids Res. 28 (2000) pp. 235–242.
- [5] J. M. Boyer and W. J. Myrvold, *On the cutting edge: simplified $O(n)$ planarity by edge addition*, J. Graph Algorith. Appl. 8 (2004) pp. 241–273.
- [6] E. Bullmore and O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems*, Nature Rev. Neurosci., 10 (2009) pp. 186–198.
- [7] A. Cardillo, S. Scellato, V. Latora, and S. Porta, *Structural properties of planar graphs of urban street patterns*, Phys. Rev. E, 73 (2006) pp. 066107.
- [8] L. F. Costa, O. N. Oliveira Jr, G. Travieso, F. A. Rodrigues, P. R. Villas Boas, L. Antiqueira, M. P. Viana, and L. E. Correa Rocha, *Analyzing and modeling real-world phenomena with complex networks: a survey of applications*, Adv. Phys., 60 (2011) pp. 329–412.
- [9] E. Estrada, *Characterization of the folding degree of proteins*, Bioinformatics, 18 (2002) pp. 697–704.
- [10] E. Estrada, *The Structure of Complex Networks. Theory and Applications*, Oxford University Press, 2011.
- [11] E. Estrada, *The communicability distance in graphs*, Lin. Alg. Appl., 436 (2012) pp. 4317–4328.
- [12] E. Estrada, *Complex networks in the Euclidean space of communicability distances*, Phys. Rev. E, 85 (2012) pp. 066122.
- [13] E. Estrada, *Graphs and Networks*, in M. Grinfeld, ed. *Mathematical Tools for Physicists*, John Wiley & Sons, 2014.
- [14] E. Estrada and N. Hatano, *Statistical-mechanical approach to subgraph centrality in complex networks*, Chem. Phys. Lett., 439 (2007) pp. 247–251.
- [15] E. Estrada and N. Hatano, *Communicability in complex networks*, Phys. Rev. E, **77** (2008) pp. 036111.
- [16] E. Estrada, N. Hatano, and M. Benzi, *The physics of communicability in complex networks*, Phys. Rep., 514 (2012) pp. 89–119.
- [17] E. Estrada and D. J. Higham, *Network properties revealed through matrix functions*, SIAM Rev., 52 (2010) pp. 696–714.
- [18] E. Estrada and J.A. Rodríguez-Velázquez, *Subgraph centrality in complex networks*, Phys. Rev. E, 71 (2005) pp. 056103.
- [19] E. Estrada, M. G. Sanchez-Lirola, and J. A. de la Peña, *Hyperspherical Embedding of Graphs and Networks in Communicability Spaces*, Discr. Appl. Math., 176 (2014) pp. 53–77.
- [20] P. J. Fleming and F. M. Richards, *Protein packing: dependence on protein size, secondary structure and amino acid composition*, J. Mol. Biol. 299 (2000) pp. 487–498.
- [21] A. Ghosh, S. Boyd, and A. Saberi, *Minimizing effective resistance of a graph*, SIAM Rev., 50 (2008) pp. 37–66.
- [22] R. L. Graham and N. J. A. Sloane, *Penny packing and two-dimensional codes*, Discr. Comput. Geom., 5 (1990) pp. 1–11.
- [23] J. L. Gross and T. W. Tucker, *Topological Graph Theory*, Dover Pub., Inc., Mineola, N.Y., 1987.
- [24] J.L. Gross, J. Yellen, and P. Zhang, eds., *Handbook of Graph Theory*, CRC Press, Boca Raton, 2013.
- [25] S. Hoory, N. Linial, and A. Wigderson, *Expander graphs and their applications*, Bull. Am. Math. Soc., 43 (2006) pp. 439–561.
- [26] B. Jiang and C. Claramunt, *Topological analysis of urban street networks*, Environ. Plan. B, 31 (2004) pp. 151–162.
- [27] D. J. Klein and M. Randić, *Resistance distance*, J. Math. Chem., 12 (1993) pp. 81–95.
- [28] A. Lubotzky, *Expander graphs in pure and applied mathematics*, Bull. Am. Math. Soc., 49 (2012) pp. 113–162.
- [29] C. R. Myers, *Software systems as complex networks: Structure, function, and evolvability of software collaboration graphs*, Phys. Rev. E, 68 (2003) pp. 046116.
- [30] M. E. J. Newman, *The structure and function of complex networks*, SIAM Rev., 45 (2003) pp. 167–256.
- [31] P. Sarnak, *WHAT IS... an Expander?*, Notices Am. Math. Soc., 51 (2004) pp. 762–763.
- [32] A. Sarzynski and A. Levy, *Spatial Efficiency and Regional Prosperity: A Literature Review and Policy Discussion*; prepared as background for GWIPP's — Implementing Regionalism project, funded by the Surdna Foundation. <http://www.gwu.edu/~gwipp/SpatialEfficiencyWPAug16.pdf> (downloaded on 20 November 2014).
- [33] L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R. F. Boyko, B. D. Sykes, and D. S. Wishart, *VADAR: a web server for quantitative evaluation of protein structure quality*, Nucleic

- Acids Res. 31 (2003) pp. 3316-3319.
- [34] W. Xiao and I. Gutman, *Resistance distance and Laplacian spectrum*, Theor. Chem. Acc., 110 (2003) pp. 284-289.